

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 15:52:14 ; Search time 170.72 Seconds

Sequence: 1 SEFDRQEYBECKRQCMQLET..... RCVSQCDKRFEEEDIDWSKYD 45
(without alignments)
17.898 Million cell updates/sec

Title: US-09-331-631A-1_COPY_29_73
Perfect score: 252
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 6790655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR:66;*

1: pir1;*
2: pir2;*
3: pir3;*
4: pir4;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	77.5	30.8	509	2 S08059	alpha-globulin type B precursor (tandem 1) - upland cotton (fragment)
2	77.5	30.8	588	1 FWNAB	N: Alternate names: seed storage protein C: Species: Gossypium hirsutum (upland cotton) C: Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 30-Sep-1993
3	72	28.6	605	2 S06398	C: Accession: S08059
4	70.5	28.0	565	2 S22477	R; Chian, C.A.; Borroto, K.; Kamalay, J.A.; Dure III, L.
5	67.5	26.8	47	2 JC3557	Plant Mol. Biol. 9, 533-546, 1987
6	67	26.6	810	2 T44430	A: Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX.
7	67	26.6	1170	1 TSHP1	A: Status: not compared with conceptual translation
8	67	26.6	1170	2 A40558	A: Molecule type: DNA
9	64	25.4	554	2 T45640	A: Residues: 1-509 <CHL>
10	63	25.0	524	2 JQ1730	C: Superfamily: glycinin
11	59.5	23.6	3078	2 T28432	Query Match 30.8%; Score 77.5; DB 2; Length 509; Best Local Similarity 41.2%; Pred. No. 0.11; Matches 14; Conservative 9; MisMatches 10; Indels 1; Gaps 1;
12	58	23.0	242	2 T29599	RESULT 2 FWNAB alpha-globulin B precursor (clone C72) - upland cotton
13	58	23.0	1421	2 T05892	N: Alternate names: seed storage protein; vicilin precursor
14	57	22.6	316	2 G71600	C: Species: Gossypium hirsutum (upland cotton)
15	57	22.6	725	1 WMYZ94	C: Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999
16	57	22.6	726	2 JQ2162	C: Accession: A30838; S06911 R; Chian, C.A.; Pyle, J.B.; Degocki, A.B.; Dure III, L.
17	57	22.6	1284	1 WMVZAT	A: Title: Developmental biochemistry of cottonseed embryogenesis and germination XVIII
18	55	21.8	305	2 B55346	A: Reference number: A30838
19	55	21.8	348	2 A71601	A: Molecule type: mRNA
20	55	21.8	623	2 T06574	A: Residues: 1-588 <CHL>
21	55	21.8	795	1 HHC108	A: Cross-references: AGBI:ML6891; NID:q167374; PID:AAA33071.1; PID:q167375
22	55	21.8	822	2 T02824	A: Experimental source: var. Coker 201
23	54.5	21.6	244	2 S44822	R; Chian, C.A.; Borroto, K.; Kamalay, J.A.; Dure III, L.
24	54.5	21.6	1021	2 T15765	Plant Mol. Biol. 9, 533-546, 1987
25	54.5	21.6	1259	2 T3201	A: Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX.
26	54.5	21.6	2715	2 T13049	A: Reference number: S06398
27	54	21.4	291	2 S62730	A: Accession: S06911
28	54	21.4	314	2 T02864	A: Status: not compared with conceptual translation
29	54	21.4	314	2 T19225	A: Molecule type: DNA
54	2	643	2	T19225	A: Residues: 1-81 <CH2>
54	2	643	2	T19225	C: Comment: This is a seed storage protein.

ALIGNMENTS

30 54 21.4 880 2 F75103
31 53.5 21.2 67 2 T15592
32 53.5 21.2 301 2 T54209
33 53 21.0 1542 2 T08880
34 53 21.0 622 2 E69006
35 53 21.0 702 2 S46854
36 53 21.0 872 2 B72167
37 53 21.0 702 2 T28570
38 53 21.0 795 2 T50255
39 52.5 20.8 91 1 S02119
40 52.5 20.8 96 1 XLHOA
41 52.5 20.8 275 2 B81430
42 52.5 20.8 425 2 T18592
43 52.5 20.8 600 2 T18593
44 52.5 20.8 1172 1 TSHP2
45 52 20.6 33 2 A41822

probable purine NTP
hypothetical prote
hypothetical prote
NMDA receptor-bind
glutamate synthase
A28L protein - var
A29L protein - var
hypothetical prote
108K heat shock pr
ubiquinol-cytochr
collipase A precurs
probable prephenat
hypothetical prote
hypothetical prote
thrombospondin 2 p
antimicrobial pept

C;Superfamily: glycinin
 C;Keywords: glycoprotein; seed; storage protein
 F;1-25;Domain: signal sequence #status predicted <SIG>
 F;26-508/Product: alpha-globulin storage proprotein #status predicted <MAT>
 F;417/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 30.8%; Score 77.5; DB 1; Length 588;
 Best Local Similarity 40.5%; Pred. No. 0.13; Mismatches 11; Indels 1; Gaps 1;
 Matches 15; Conservative 9; Mismatches 12; Indels 1; Gaps 1;

Qy 2 EFDQEVEECKRQCMOLETS-GOMRCVSCDKRFEED 38
 Db 81 RDPQRYYEECQQECRQOE-ERORPQQCQRCLKRFEE 116

RESULT 3
 S06398
 alpha-globulin type A precursor - upland cotton
 N;Alternate names: seed storage protein
 C;Species: Gossypium hirsutum (upland cotton)
 C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 30-Sep-1993
 R;Chian, C.A.; Borroto, K.; Kamalay, J.A.; Dure III, L.
 Plant Mol. Biol. 9, 533-546, 1987
 A;Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX. Se
 A;Reference number: S06398
 A;Accession: S06398
 A;Status: not compared with conceptual translation
 A;Molecule type: DNA
 A;Residues: 1-1605 <CHI>
 C;Superfamily: glycinin
 F;1-24/Domain: signal sequence #status predicted <SIG>
 F;25-505/Product: alpha-globulin type A #status predicted <MAT>

RESULT 5
 JC557
 arginine/glutamate-rich 6.5K polypeptide - smooth loofah
 C;Species: Luffa cylindrica (smooth loofah)
 C;Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 07-May-1999
 C;Accession: JC557
 C;Reference number: R;Kimura, M.; Park, S.S.; Sakai, R.; Yamasaki, N.; Funatsu, G.
 Blosci, Biotechnol. Biochem. 61, 984-988, 1997
 A;Title: Primary structure of 6.5k-arginine/glutamate-rich polypeptide from the seeds
 A;Accession number: JC557; MUID:9735743
 A;Molecule type: protein
 A;Residues: 1-47 <KIM>
 A;Experimental source: seed
 C;Comment: This protein is a storage protein which provides nitrogen and carbon reser
 F;12-33,16-29/Disulfide bonds: #status predicted

Query Match 28.6%; Score 72; DB 2; Length 605;
 Best Local Similarity 42.5%; Pred. No. 0.57; Mismatches 17; Indels 4; Gaps 3;
 Matches 17; Conservative 9; Mismatches 10; Indels 4; Gaps 3;

Qy 1 SEFD-ROEYEECKRQCMOLETS-GOMRCVSCDKRFEEDI 39
 Db 33 SDPPQDQRYEDCRKRC-QLETRQTEQ--DRCEDRSSETOL 69

RESULT 4
 S22477
 vicilin precursor - cacao
 C;Species: Theobroma cacao (cacao)
 C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Mar-1999
 C;Accession: S22477; S22478; S18105; S22050
 R;McHenry, L.; Fritz, P.J.
 Plant Mol. Biol. 18, 1173-1176, 1992
 A;Title: Comparison of the structure and nucleotide sequence of vicilin genes of cocoa
 A;Reference number: S22477; MUID:92288309
 A;Accession: S22477
 A;Molecule type: DNA
 A;Residues: 1-1566 <MCB>
 A;Cross-references: EMBL:X62625
 A;Accession: S22478
 A;Molecule type: mRNA
 A;Residues: 1-810 <YAM>
 A;Cross-references: EMBL:AB019195; NID:93808061; PIDN:BA34056.1; PID:93808062

RESULT 6
 T44430
 protein Pv100 [Imported] - winter squash
 C;Species: Cucurbita maxima (winter squash)
 C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
 C;Accession: T44430
 R;Yamada, K.; Shimada, T.; Kondo, M.; Nishimura, M.; Hara-Nishimura, I.
 J. Biol. Chem. 274, 2563-2570, 1999
 A;Title: Multiple functional proteins are produced by cleaving Asn-Gln bonds of a sin
 A;Reference number: 222767; MUID:99107919
 A;Accession: T44430
 A;Molecule type: translated from GB/EMBL/DDBJ
 A;Residues: 1-810 <YAM>
 A;Cross-references: EMBL:AB019195; NID:93808061; PIDN:BA34056.1; PID:93808062

Query Match 26.6%; Score 67; DB 2; Length 810;
 Best Local Similarity 34.1%; Pred. No. 2.8; Mismatches 15; Indels 8; Gaps 2;
 Matches 15; Conservative 6; Mismatches 15; Indels 8; Gaps 2;

Qy 5 ROEYEECKRQCMOLETS-GOMRCVSCDKRFE----- 40
 Db 75 RAEEYEVCRRLCQVAERGVEQQRKEQVCEERLREREQGRGEDVD 118

RESULT 7
 TSHP1
 thrombospondin 1 precursor - human
 C;Species: Homo sapiens (man)
 C;Date: 23-Aug-1987 #sequence_revision 03-Aug-1995 #text_change 13-Aug-1999
 C;Accession: A26155; A34274; A30140; A25612; A05172; A42927
 R;Lawler, J.; Hynes, R.O.
 J. Cell Biol. 103, 1635-1648, 1986
 A;Title: The structure of human thrombospondin, an adhesive glycoprotein with multipl
 A;Reference number: A26155; MUID:87057617
 A;Accession: A26155
 A;Molecule type: mRNA
 A;Residues: 1-1170 <LAW>

Query Match 28.0%; Score 70.5; DB 2; Length 566;
 Best Local Similarity 32.4%; Pred. No. 0.8; Mismatches 11; Indels 1; Gaps 1;
 Matches 11; Conservative 11; Mismatches 11; Indels 1; Gaps 1;

A;Cross-references: GB:X04665; NID:q3137; PIDN:CAA8370_1; PID:g31138
A;Note: parts of this sequence, including the amino end of the mature protein, were determined by sequencing of the cDNA clone.
R;Laherty, C.D.; Gierman, T.M.; Dixit, V.M.
J. Biol. Chem. 264, 11222-11227, 1989

A;Title: Characterization of the promoter region of the human thrombospondin gene. DNA sequence analysis and transcriptional activity. J. Cell Biol. 108, 729-736, 1989

A;Accession: A30140
A;Reference number: A30140; MUID:89139590

A;Molecule type: mRNA
A;Cross references: EMBL:XI4787; NID:937464; PIDN:CAA32899_1; PID:g37465
A;Note: parts of this sequence, including the amino end of the mature protein, were determined by analysis of the cDNA clone.
R;Kobayashi, S.; Eden-McCutchan, F.; Framson, P.; Bernstein, P.
Biochemistry 25, 8418-8425, 1986

A;Title: Partial amino acid sequence of human thrombospondin as determined by analysis of the cDNA clone. Reference number: A25812; MUID:87157592

A;Accession: A25812
A;Molecule type: mRNA
A;Residues: 1-83, 'A', 85-397 <KOB>
A;Cross references: GB:M25631; NID:9538353; PIDN:AAA36741_1; PID:g538354
R;Dixit, V.M.; Hennessy, S.W.; Grant, G.A.; Rotwein, P.; Bernstein, P.
Proc. Natl. Acad. Sci. U.S.A. 83, 5449-5453, 1986

A;Reference number: A05172; MUID:86287276

A;Accession: A05172
A;Molecule type: mRNA
A;Residues: 1-83, 'A', 85-397 <LAW>
A;Cross references: GB:M2465; NID:9538353; PIDN:AAA36741_1; PID:g538354
R;Sun, X.; Skorstengaard, K.; Mosher, D.F.
J. Cell Biol. 118, 693-701, 1992

A;Title: Disulfides modulate RGD-inhibitable cell adhesive activity of thrombospondin. Reference number: A42927; MUID:92348511

A;Accession: A42927
A;Molecule type: protein
A;Residues: 987-1003 <SUN>
A;Note: Cys-992 is shown to have a free sulphydryl
A;Genetics:
A;Gene: GDB:THBS1; TSP
A;Cross-references: GDB:120438; OMIM:188050
A;Map position: 15q15-15q15
A;Introns: 23/1
A;Note: the list of introns may be incomplete
C;Function:
C;Description: participates in cell migration and adhesion, and in platelet aggregation
C;Superfamily: thrombospondin type I repeat homology; von Willebrand factor type C repeat homology; vWF
C;Keywords: beta hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trimeric F1-18/Domain: signal sequence #status predicted <SIG>
F1-19-170/Domain: thrombospondin type I repeat homology <MAT>
F1-17-375/Domain: von Willebrand factor type C repeat homology <VWCF>
F1-18/Domain: signal sequence #status predicted <SIG>
F1-19-170/Domain: thrombospondin type I repeat homology <MAT>
F317-375/Domain: von Willebrand factor type C repeat homology <VWCF>
F378-429/Domain: thrombospondin type I repeat homology <THR2>
F491-547/Domain: thrombospondin type I repeat homology <THR2>
F551-586/Domain: EGF homology <EGF>
F65-689/Domain: EGF homology <EGF>
F92-928/Region: cell attachment (R-G-D) motif
F171-232/disulfide bonds: #status predicted
F248-360,708,1067/site: carbohydrate (Asn) (covalent) #status predicted
F270-274/disulfide bonds: interchain #status predicted
F610/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
F1051/Binding site: carbohydrate (Asn) (covalent) #status absent

Query Match 26.6%; Score 67; DB 1; Length 1170;
Best Local Similarity 39.5%; Pred. No. 3.9; Length 1170;
Matches 15; Conservative 6; Mismatches 11; Indels 6; Gaps 2;

RESULT 9
#45840
RSCDSLNRCGGSSVQRTCHIQECDKRFKQDGWSH 441

QY 13 RQOMOLE---TSGQMRCC-VSQCDKRFEDIDWSKY 44
| | | | : | | | | : || || | : || | | : || | |
Db 404 RSCDSLNRCGGSSVQRTCHIQECDKRFKQDGWSH 441

R;Rawler, J.; Duquette, M.; Ferro, P.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.
Genomics 1, 587-600, 1991

A;Title: Characterization of the murine thrombospondin gene. Reference number: A40558; MUID:92128941

A;Accession: A40558
A;Status: preliminary

A;Molecule type: DNA
A;Residues: 1-1170 <LAW>
A;Cross references: GB:M62449; GB:M62450; GB:M62451; GB:M62452; GB:M62453; GB:M62454; M62465; GB:M2466; GB:M62467; GB:M62468; GB:M62469; GB:M2470; NID:9511867; PIDN:AAA57905
R;Bernstein, P.; Alfie, D.; Devavarayulu, S.; Framson, P.; Li, P.
J. Biol. Chem. 265, 16691-16698, 1990

A;Title: Characterization of the mouse thrombospondin gene and evaluation of the role of the mouse thrombospondin gene. Reference number: A37905; MUID:90375546

A;Accession: A37905
A;Status: preliminary

A;Molecule type: DNA
A;Residues: 1-490 <BOR>
A;Cross references: GB:J05605; GB:J05606; NID:9201991; PIDN:AAA40431_1; PID:g554390
R;Laherty, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seloin, M.F.; Dixit, V.M.
J. Biol. Chem. 267, 33274-33281, 1992

A;Title: Characterization of mouse thrombospondin 2 sequence and expression during development. Reference number: A42587; MUID:92147683

A;Accession: B42587
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1152, 'P', 1154-1170 <LAH>
A;Cross references: GB:M8276
A;Title: Sequence extracted from NCBI backbone (NCBIPR:81501)
R;Chen, H.; Aeschlimann, D.; Nowlen, J.; Mosher, D.F.
FEBS Lett. 387, 36-41, 1996

A;Title: Expression and initial characterization of recombinant mouse thrombospondin 2. Reference number: S68787; MUID:96234006

A;Accession: S68787
A;Molecule type: protein
A;Residues: 19-26, 'X', 28-37 <CHE>
C;Complex: homotrimer; disulfide linked
C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology;
C;Keywords: calcium binding; glycoprotein; homotrimer
F1-1-18/Domain: signal sequence #status predicted <SIG>
F1-19-170/Domain: thrombospondin type I repeat homology <MAT>
F1-17-375/Domain: von Willebrand factor type C repeat homology <VWCF>
F1-18/Domain: signal sequence #status predicted <SIG>
F1-19-170/Domain: thrombospondin type I repeat homology <MAT>
F317-375/Domain: von Willebrand factor type C repeat homology <VWCF>
F378-429/Domain: thrombospondin type I repeat homology <THR2>
F491-547/Domain: thrombospondin type I repeat homology <THR2>
F551-586/Domain: EGF homology <EGF>
F248,360,708,1067/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 26.6%; Score 67; DB 2; Length 1170;
Best Local Similarity 39.5%; Pred. No. 3.9; Length 1170;
Matches 15; Conservative 6; Mismatches 11; Indels 6; Gaps 2;

RESULT	14	Db	550 KRRNVEWLS-RLRRDIKECDK-YKEDLDKA 579
rifin	PFB1005w - malaria parasite (Plasmodium falciparum)	Search completed:	March 1, 2001, 15:52:17
C;Species:	Plasmodium falciparum	Job time:	541 sec
C;Date:	13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000		
R;Gardner, M.J.; Nettelein, H.; Carrucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;			
; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.			
A;Title:	Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.		
A;Reference number:	A71600; MUID:99021743		
A;Accession:	G71600		
A;Status:	preliminary; nucleic acid sequence not shown; translation not shown		
A;Molecule type:	DNA		
A;Residues:	1-316 <GAR>		
A;Cross-references:	GB:AE001432; GB:AE001362; NID:93845332; PIDN: AAC71988.1; PID:9384533		
A;Experimental source:	clone 3D7		
A;Genes:	PFB1005w		
RESULT	15		
WW7294	Query Match 22.6%; Score 57; DB 2; Length 316; Best Local Similarity 34.9%; Pred. No. 18; Matches 15; Conservative 4; Mismatches 10; Indels 14; Gaps 2;		
Qy	3 FDHQ-----EYERCKRQCMOLETSQGMRRCYVSQCDRFEEDI 39		
Db	67 FDHQTSERFEEDEE-----RMKDKKRKCEQQCDKQEI 101		
RESULT	15		
WW7294	Query Match 22.6%; Score 57; DB 2; Length 316; Best Local Similarity 34.9%; Pred. No. 18; Matches 15; Conservative 4; Mismatches 10; Indels 14; Gaps 2;		
C;Species:	vaccinia virus		
A;Note:	host Homo sapiens (man)		
C;Date:	31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 29-Oct-1999		
C;Accession:	A41701; A40825; S29908		
R;De Carlos, A.; Paez, E.			
Virology 185, 768-778, 1991			
A;Title:	Isolation and characterization of mutants of vaccinia virus with a modified A-type inclusion protein.		
A;Reference number:	A41701; MUID:92074241		
A;Accession:	A41701		
A;Molecule type:	DNA		
A;Residues:	1-725 <DEC>		
A;Cross-references:	GB:M76371; NID:9335683; PIDN:AAA48275.1; PID:9335684		
R;Amegadze, B.Y.; Sisler, J.R.; Moss, B.			
Virology 186, 777-782, 1992			
A;Title:	Frame-shift mutations within the vaccinia virus A-type inclusion protein gene.		
A;Reference number:	A40825; MUID:92124754		
A;Accession:	A40825		
A;Molecule type:	DNA		
A;Residues:	1-586, 'KO', 589-609, 'R', 611-618; 620-682, 'S', 684-725 <AME>		
A;Cross-references:	GB:M61187; NID:9335782; PIDN:AAA48321.1; PID:9335784		
R;Amegadze, B.Y.			
submitted to the EMBL Data Library, January 1991			
A;Reference number:	S29907		
A;Accession:	S29908		
A;Status:	preliminary		
A;Molecule type:	DNA		
A;Residues:	1-586, 'KO', 589-609, 'R', 611-618; 620-682, 'S', 684-725 <AM2>		
A;Cross-references:	EMBL:X57318; NID:962249; PIDN:CAA46574.1; PID:962241		
C;Superfamily:	covox virus A-type inclusion protein		
C;Keywords:	inclusion protein		
Query Match 22.6%; Score 57; DB 1; Length 725; Best Local Similarity 40.6%; Pred. No. 38; Matches 13; Conservative 11; Mismatches 6; Indels 2; Gaps 2;			
Qy	12 KRCQMQLETSQGMRRCYVSQCDKRFEDIDWSK 43		

